

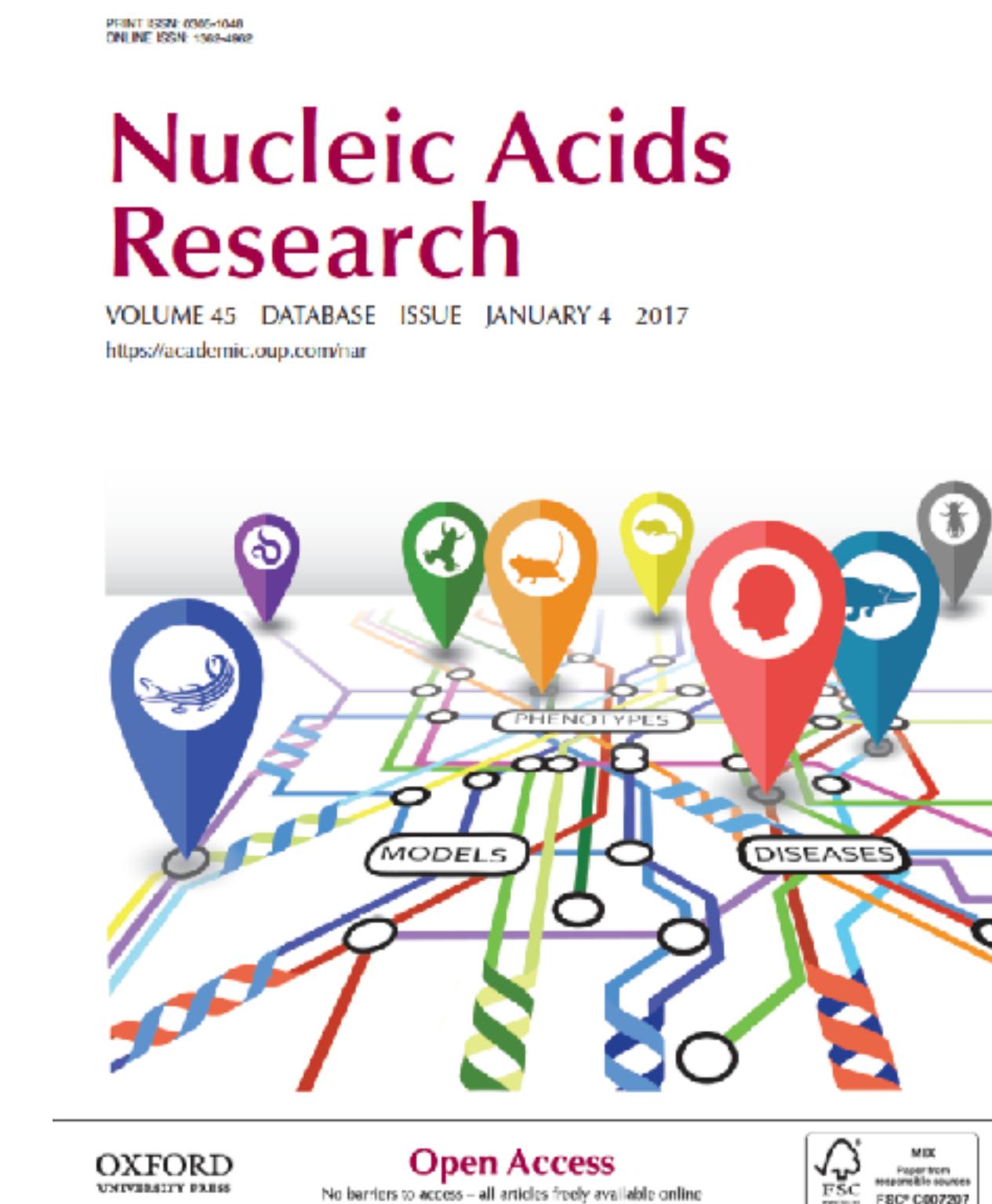
An overview of bioinformatics databases and online resources: what they are and how to access them

Mark Stenglein, GDW



There are an overwhelming number of databases and other online resources, which often have overlapping content and purpose

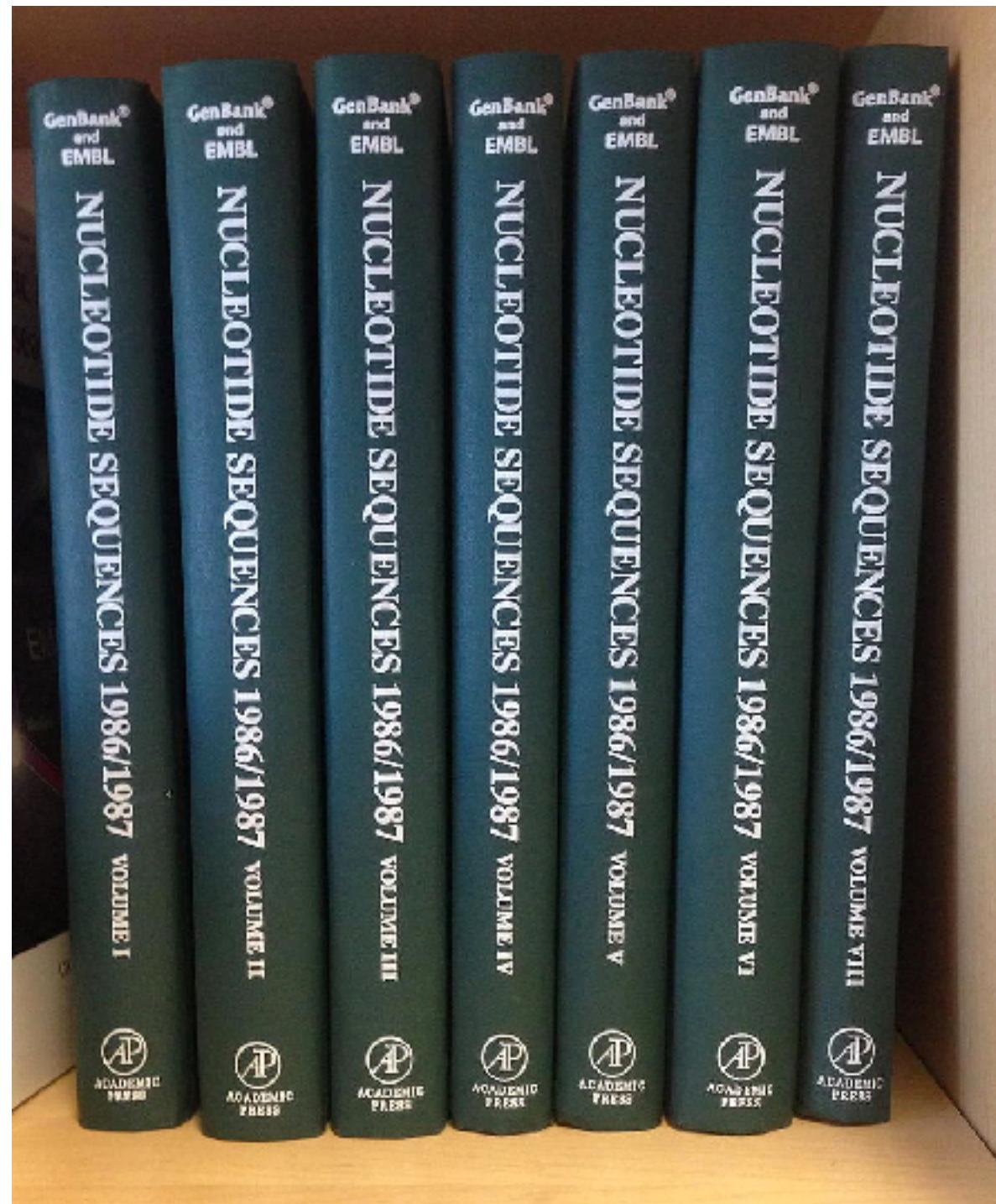
The annual Database and Web Server NAR issue is a good resource



<https://academic.oup.com/nar/issue/45/D1>

GenBank was one of the earliest sequence databases.

GenBank circa 1987



~10,000 sequences

GenBank release 100 (1997) distributed by CDROM



~1,300,000 sequences

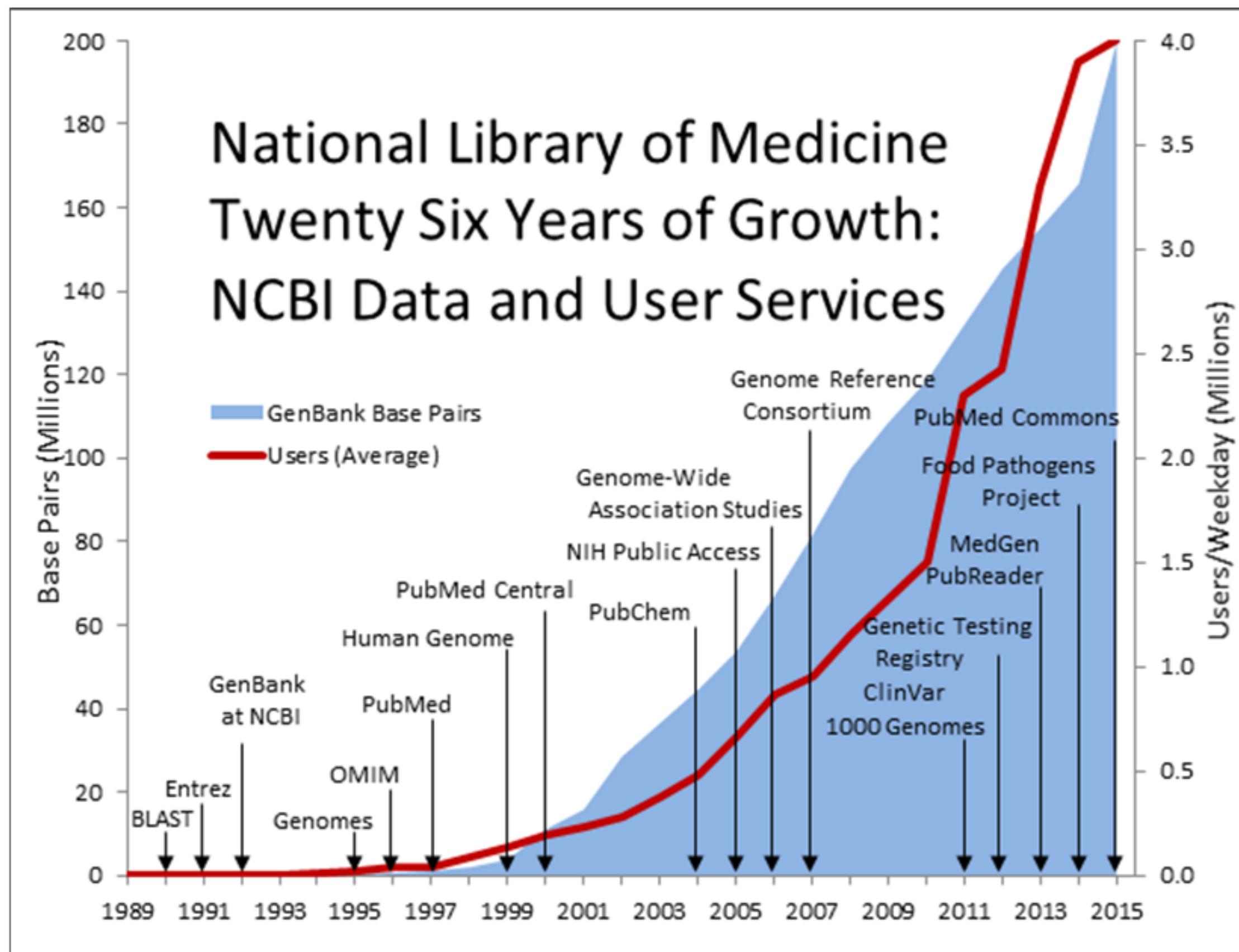
Genbank today



>200,000,000 sequences

Today, we'll focus mainly on NCBI databases and resources, and how to access them

The NCBI was created in 1987 by the US government



Categories of NCBI databases

Category	Example NCBI db	Content
Literature	PubMed	Scientific and medical abstracts/citations
Genomes	Assembly	Genome assembly information
Genes	Gene	Collected information about gene loci
Proteins	Protein	Protein sequences
Chemicals	PubChem Compound	Chemical information with structures, information and links
Health	dbGaP	Genotype/phenotype interaction studies

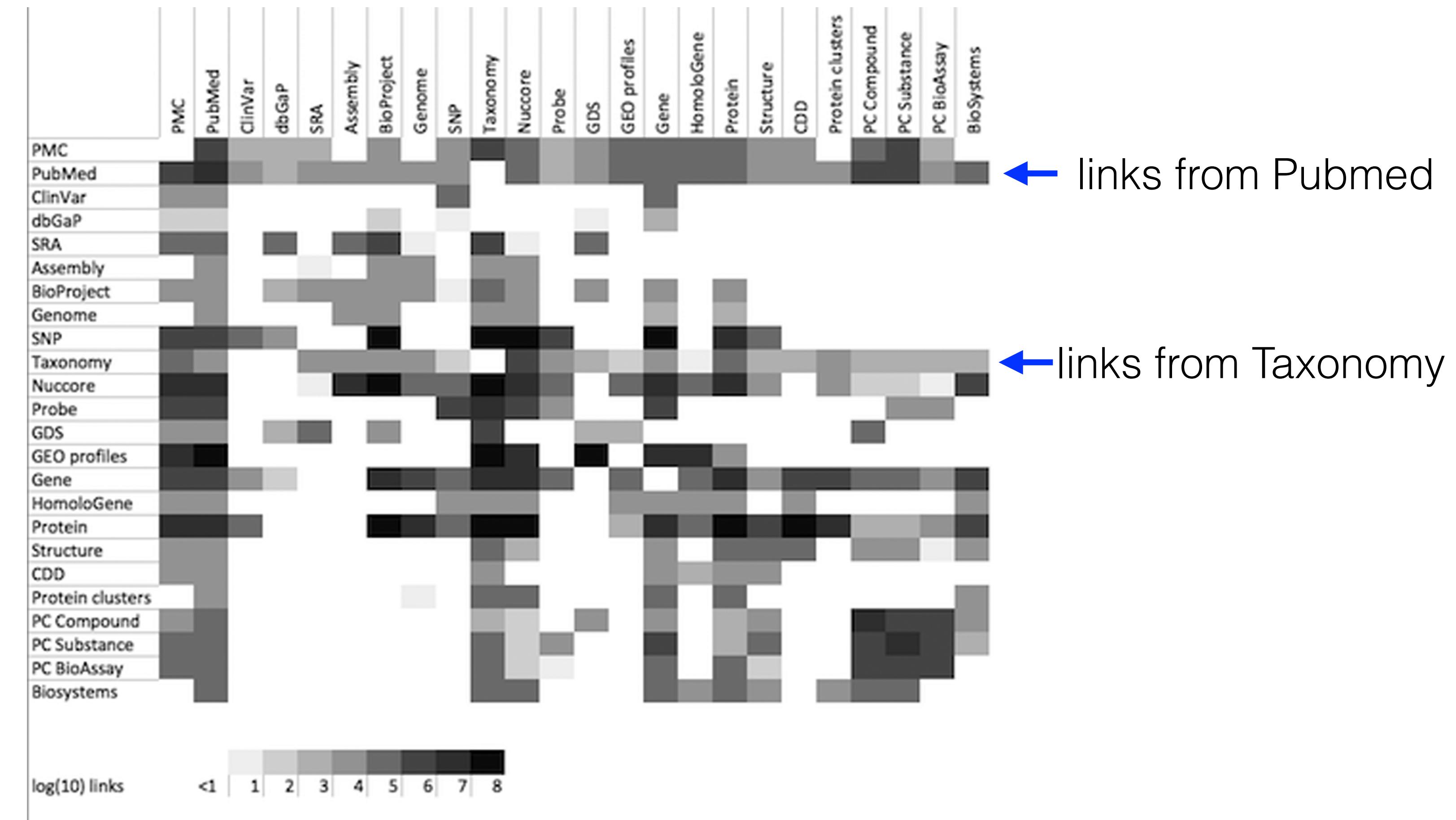
image: NIH/NLM

<https://academic.oup.com/nar/issue/45/D1>

One really useful feature of NCBI databases is that they link to each other

So, you can, for example:

- get all the nucleotide sequences associated with a taxon of interested
- get all the protein sequences predicted to be encoded by a genome
- get the SRA datasets associated with a particular paper in Pubmed



Say we were interested in nucleotide sequences from Bob's papers

The screenshot shows the PubMed search results for the query "fitak RR". A blue arrow points to the search term "fitak RR" in the search bar. The results page displays 19 items, sorted by "Most Recent".

Search results: Items: 19

- The Expectations and Challenges of Wildlife Disease Research in the Era of Genomics: Forecasting with a Horizon Scan-like Exercise.**
Fitak RR, Antonides JD, Baitchman EJ, Bonacorso E, Braun J, Kubiski S, Chiu E, Fagre AC, Gagne RB, Lee JS, Malmberg JL, Stenglein MD, Dusek RJ, Forgacs D, Fountain-Jones NM, Gilbertson MLJ, Worsley-Tonks KEL, Funk WC, Trumbo DR, Ghersi BM, Grimaldi W, Heisel SE, Jardine CM, Kamath PL, Karmacharya D, Kozakiewicz CP, Kraberger S, Loisel DA, McDonald C, Miller S, O'Rourke D, Ott-Conn CN, Páez-Vacas M, Peel AJ, Turner WC, VanAcker MC, VandeWoude S, Pecon-Slattery J.
J Hered. 2019 May 7;110(3):261-274. doi: 10.1093/jhered/esz001.
PMID: 31067326
[Similar articles](#)
- Genomic signatures of G-protein-coupled receptor expansions reveal functional transitions in the evolution of cephalopod signal transduction.**
Ritschard EA, Fitak RR, Simakov O, Johnsen S.
Proc Biol Sci. 2019 Feb 27;286(1897):20182929. doi: 10.1098/rspb.2018.2929.
PMID: 30963849 [Free PMC Article](#)
[Similar articles](#)
- Spectral sensitivity in ray-finned fishes: diversity, ecology and shared descent.**
Schweikert LE, Fitak RR, Caves EM, Sutton TT, Johnsen S.
J Exp Biol. 2018 Nov 27;221(Pt 23). pii: jeb189761. doi: 10.1242/jeb.189761.
PMID: 30322978 [Free Article](#)
[Similar articles](#)
- Near absence of differential gene expression in the retina of rainbow trout after exposure to a magnetic pulse: implications for magnetoreception.**
Fitak RR, Schweikert LE, Wheeler BR, Ernst DA, Lohmann KJ, Johnsen S.
Biol Lett. 2018 Jun;14(6). pii: 20180209. doi: 10.1098/rsbl.2018.0209.
PMID: 29875210
[Similar articles](#)
- Genome-Wide Analysis of SNPs Is Consistent with No Domestic Dog Ancestry in the Endangered Mexican Wolf (*Canis lupus baileyi*).**
Fitak RR, Rinkevich SE, Culver M.
J Hered. 2018 May 11;109(4):372-383. doi: 10.1093/jhered/esy009.
PMID: 29757430 [Free PMC Article](#)
[Similar articles](#)

Filters: Manage Filters

Sort by: Best match (highlighted) Most recent

Find related data

Database: Select

Search details: fitak RR[Author]

Recent Activity:

- Turn Off Clear
- fitak RR (19)
- fitak R (20)
- fitak B (20)
- Wolbachia-mediated virus blocking in mosquito cells is dependent on XRN1-
- A genome-wide RNAi screen reveals that mRNA decapping restricts bunyaviral repl...

See more...

Get nucleotide sequences associated with these papers

We've moved to a new database

NCBI Resources How To Sign in to NCBI

Nucleotide Nucleotide Advanced Search Help

The Nucleotide database will include EST and GSS sequences in early 2019. [Read more.](#)

Species Summary 20 per page Sort by Default order Send to: Filters: [Manage Filters](#)

Animals (67) Customize ...

Molecule types genomic DNA/RNA (66) Customize ...

Source databases INSDC (GenBank) (67) Customize ...

Sequence Type Nucleotide (67)

Genetic compartments Mitochondrion (40)

Sequence length Custom range...

Release date Custom range...

Revision date Custom range...

[Clear all](#)

[Show additional filters](#)

Items: 1 to 20 of 67 << First < Prev Page 1 of 4 Next > Last >>

[Sorex monticolus isolate RF-3 cytochrome b \(cytb\).gene, complete cds; mitochondrial](#)
1. 1,140 bp linear DNA
Accession: JF489124.1 GI: 326632150
[Prtein](#) [PubMed](#) [Taxonomy](#)
[GenBank](#) [FASTA](#) [Graphics](#) [PopSet](#)

[Lepus insularis isolate RF-3 cytochrome b \(cytb\).gene, complete cds; mitochondrial](#)
2. 1,140 bp linear DNA
Accession: HM222713.1 GI: 298162478
[Prtein](#) [PubMed](#) [Taxonomy](#)
[GenBank](#) [FASTA](#) [Graphics](#) [PopSet](#)

[Phocoena sinus isolate RF-9 cytochrome b \(cytb\).gene, complete cds; mitochondrial](#)
3. 1,140 bp linear DNA
Accession: HM222714.1 GI: 298162480
[Prtein](#) [PubMed](#) [Taxonomy](#)
[GenBank](#) [FASTA](#) [Graphics](#) [PopSet](#)

[Didelphis virginiana isolate RF-10 cytochrome b \(cytb\).gene, complete cds; mitochondrial](#)
4. 1,149 bp linear DNA
Accession: HM222715.1 GI: 298162482
[Prtein](#) [PubMed](#) [Taxonomy](#)
[GenBank](#) [FASTA](#) [Graphics](#) [PopSet](#)

[Canis lupus familiaris cytochrome b \(cytb\).gene, complete cds; mitochondrial](#)
5. 1,140 bp linear DNA
Accession: JF489119.1 GI: 326632140
[Prtein](#) [PubMed](#) [Taxonomy](#)
[GenBank](#) [FASTA](#) [Graphics](#) [PopSet](#)

[Trichechus manatus cytochrome b \(cytb\).gene, complete cds; mitochondria](#)
6. 1,140 bp linear DNA
Accession: JF489120.1 GI: 326632142

Results by taxon Top Organisms [Tree]
Camelus bactrianus (10)
Camelus dromedarius (9)
Camelus ferus (7)
Cervus elaphus (1)
Alces alces (1)
All other taxa (39)
[More...](#)

Find related data Database: [Select](#)

Find items

Recent activity Turn Off Clear

fitak RR (19) PubMed

fitak R (20) PubMed

fitak B (20) PubMed

Wolbachia-mediated virus blocking in mosquito cells is dependent on XRN1- See more...

A genome-wide RNAi screen reveals that mRNA decapping restricts bunyaviral repli...

See more...

We're not in pubmed anymore



You could click on these sequences one at a time

The screenshot shows the NCBI Nucleotide database interface. The search bar at the top contains the query "Nucleotide". Below the search bar, there is a message: "The Nucleotide database will include EST and GSS sequences in early 2019. [Read more.](#)". The main content area displays the details for the sequence JF489124.1, which is the GenBank entry for the *Sorex monticolus* cytochrome b (cytb) gene, complete cds; mitochondrial. The sequence information includes:

- LOCUS**: JF489124
- DEFINITION**: *Sorex monticolus* cytochrome b (cytb) gene, complete cds; mitochondrial.
- ACCESSION**: JF489124
- VERSION**: JF489124.1
- KEYWORDS**: .
- SOURCE**: mitochondrion *Sorex monticolus* (montane shrew)
- ORGANISM**: *Sorex monticolus*
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Insectivora; Soricidae;
Soricinae; Sorex.
- REFERENCE**: 1 (bases 1 to 1140)
AUTHORS: Naidu,A., Fitak,R.R., Munguia-Vega,A. and Culver,M.
TITLE: Novel primers for complete mitochondrial cytochrome b gene sequencing in mammals
JOURNAL: Mol Ecol Resour 12 (2), 191-196 (2012)
PUBMED: [21974833](#)
REFERENCE: 2 (bases 1 to 1140)

On the right side of the page, there are several interactive buttons and dropdown menus:

- Send to:** Change region shown, Customize view
- Analyze this sequence:** Run BLAST, Pick Primers, Highlight Sequence Features, Find in this Sequence
- Related information:** Protein, PubMed, Taxonomy, PopSet

Or you can download them all at once, in various formats

NCBI Resources How To Sign in to NCBI

Nucleotide Nucleotide Advanced Search Help

The Nucleotide database will include EST and GSS sequences in early 2019. [Read more.](#)

Species Summary ▾ 20 per page ▾ Sort by Default order ▾ Send to: ▾ Filters: [Manage Filters](#)

Animals (67) Animals (67)

Customize ... Customize ...

Molecule types Summary ▾ 20 per page ▾ Sort by Default order ▾ Send to: ▾ Filters: [Manage Filters](#)

genomic DNA/RNA (66) Genomic DNA/RNA (66)

Customize ... Customize ...

Source databases Summary ▾ 20 per page ▾ Sort by Default order ▾ Send to: ▾ Filters: [Manage Filters](#)

INSDC (GenBank) (67) INSDC (GenBank) (67)

Customize ... Customize ...

Sequence Type Summary ▾ 20 per page ▾ Sort by Default order ▾ Send to: ▾ Filters: [Manage Filters](#)

Nucleotide (67) Nucleotide (67)

Genetic compartments Summary ▾ 20 per page ▾ Sort by Default order ▾ Send to: ▾ Filters: [Manage Filters](#)

Mitochondrion (40) Mitochondrion (40)

Sequence length Summary ▾ 20 per page ▾ Sort by Default order ▾ Send to: ▾ Filters: [Manage Filters](#)

Custom range... Custom range...

Release date Summary ▾ 20 per page ▾ Sort by Default order ▾ Send to: ▾ Filters: [Manage Filters](#)

Custom range... Custom range...

Revision date Summary ▾ 20 per page ▾ Sort by Default order ▾ Send to: ▾ Filters: [Manage Filters](#)

Custom range... Custom range...

Items: 1 to 20 of 67

<< First < Prev Page 1 of 4 Next > Last >>

[Sorex monticolus cytochrome b \(cytb\) gene, complete cds; mitochondrial](#)

1. 1,140 bp linear DNA
Accession: JF489124.1 GI: 326632150
[Protein](#) [PubMed](#) [Taxonomy](#)
[GenBank](#) [FASTA](#) [Graphics](#) [PopSet](#)

[Lepus insularis isolate RF-8 cytochrome b \(cytb\) gene, complete cds; mitochondrial](#)

2. 1,140 bp linear DNA
Accession: HM222713.1 GI: 298162478
[Protein](#) [PubMed](#) [Taxonomy](#)
[GenBank](#) [FASTA](#) [Graphics](#) [PopSet](#)

[Phocoena sinus isolate RF-9 cytochrome b \(cytb\) gene, complete cds; mitochondrial](#)

3. 1,140 bp linear DNA
Accession: HM222714.1 GI: 298162480
[Protein](#) [PubMed](#) [Taxonomy](#)
[GenBank](#) [FASTA](#) [Graphics](#) [PopSet](#)

[Didelphis virginiana isolate RF-10 cytochrome b \(cytb\) gene, complete cds; mitochondrial](#)

Results by taxon

Top Organisms [Tree]

- Camelus bactrianus (10)
- Camelus dromedarius (9)
- Camelus ferus (7)
- Cervus elaphus (1)
- Alces alces (1)
- All other taxa (39)

More...

Find related data

Database: Select

Find items

Recent activity

Turn Off Clear

fitak RR (19)

PubMed

Or you can download them all at once, in various formats

The screenshot shows the NCBI Nucleotide database search results page. The search bar at the top contains "Nucleotide". The main content area displays "Items: 1 to 20 of 67" results, each with a checkbox, a title, an accession number, and links to Protein, PubMed, and Taxonomy. Below the results are links to GenBank, FASTA, Graphics, and PopSet. A context menu is open over the third result, listing download options: Complete Record, Coding Sequences, Gene Features, Choose Destination (File, Clipboard, Collections), and a list of formats including Summary, GenBank, GenBank (full), **FASTA** (highlighted with a blue arrow), ASN.1, XML, INSDSeq XML, TinySeq XML, Feature Table, Accession List, GI List, and GFF3. The menu also includes "Send to:" and "Filters: Manage Filters" buttons.

NCBI Resources ▾ How To ▾ Sign in to NCBI

Nucleotide Nucleotide Advanced Search Help

The Nucleotide database will include EST and GSS sequences in early 2019. [Read more.](#)

Species Summary ▾ 20 per page ▾ Sort by Default order ▾

Animals (67)
Customize ...

Molecule types genomic DNA/RNA (66)
Customize ...

Source databases INSDC (GenBank) (67)
Customize ...

Sequence Type Nucleotide (67)

Genetic compartments Mitochondrion (40)

Sequence length Custom range...

Release date Custom range...

Revision date Custom range...

Items: 1 to 20 of 67

<< First < Prev > Next >> Last

1. [Sorex monticolus cytochrome b \(cytb\) gene, complete cds; mitochondria](#)
Accession: JF489124.1 GI: 326632150
[Protein](#) [PubMed](#) [Taxonomy](#)
[GenBank](#) [FASTA](#) [Graphics](#) [PopSet](#)

2. [Lepus insularis isolate RF-8 cytochrome b \(cytb\) gene, complete cds; mitochondria](#)
Accession: HM222713.1 GI: 298162478
[Protein](#) [PubMed](#) [Taxonomy](#)
[GenBank](#) [FASTA](#) [Graphics](#) [PopSet](#)

3. [Phocoena sinus isolate RF-9 cytochrome b \(cytb\) gene, complete cds; mitochondria](#)
Accession: HM222714.1 GI: 298162480
[Protein](#) [PubMed](#) [Taxonomy](#)
[GenBank](#) [FASTA](#) [Graphics](#) [PopSet](#)

Send to: Filters: [Manage Filters](#)

Complete Record
Coding Sequences
Gene Features

Choose Destination

File Clipboard
Collections

Download 67 items.

Format

✓ Summary
GenBank
GenBank (full)
FASTA
ASN.1
XML
INSDSeq XML
TinySeq XML
Feature Table
Accession List
GI List
GFF3

[Tree]
rianus (10)
nedarius (9)
i (7)
is (1)
j (39)

Recent activity

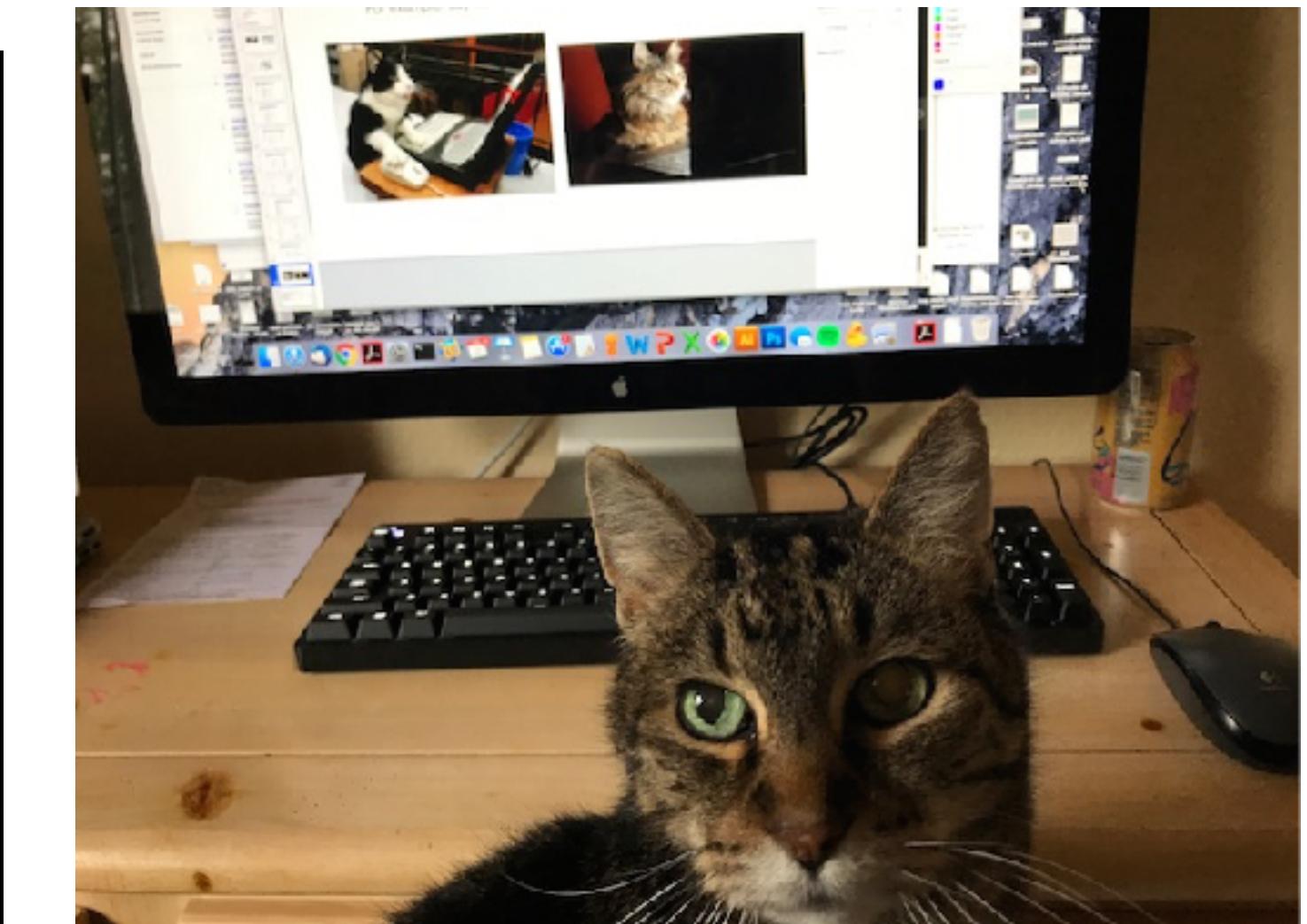
fitak RR (19)

Turn Off Clear

PubMed

There are often many paths to the same data

For example, say we want to download the cat (*Felis catus*) genome



Kirby, 17 year old male cat

You could try to get the cat genome from the NCBI nucleotide db

<https://www.ncbi.nlm.nih.gov/nuccore/?term=Felis+catus>

NCBI Resources ▾ How To ▾ Sign in to NCBI

Nucleotide Nucleotide Felis catus Create alert Advanced Help

Species Summary ▾ 20 per page ▾ Sort by Default order ▾ Send: ▾ Filters: Manage Filters

Animals (355,888)
Plants (10)
Fungi (195)
Protists (38,091)
Bacteria (5,536)
Viruses (3,369)
Customize ...

Items: 1 to 20 of 403089 ←

Found 407775 nucleotide sequences. Nucleotide (403089) EST (1089) GSS (3597)

[Felis catus polycystic kidney disease 1-like \(PKD1\) gene, exons 21, 22, 23, 24, 29, 37, and 38](#)
1. 3,269 bp linear DNA
Accession: AH014595.2 GI: 1059791198
[GenBank](#) [FASTA](#) [Graphics](#)

[Felis catus isolate Cinnamon breed Abyssinian chromosome A1, Felis_catus_8.0, whole genome](#)
2. shotgun sequence
240,380,223 bp linear DNA
Accession: NC_018723.2 GI: 753572113
[GenBank](#) [FASTA](#) [Graphics](#)

[Felis catus isolate Cinnamon breed Abyssinian chromosome A2, Felis_catus_8.0, whole genome](#)
3. shotgun sequence
168,638,799 bp linear DNA
Accession: NC_018724.2 GI: 753572104
[GenBank](#) [FASTA](#) [Graphics](#)

[Felis catus isolate Cinnamon breed Abyssinian chromosome A3, Felis_catus_8.0, whole genome](#)
4. shotgun sequence
140,925,898 bp linear DNA
Accession: NC_018725.2 GI: 753572100
[GenBank](#) [FASTA](#) [Graphics](#)

Results by taxon

Top Organisms [Tree]
Felis catus (355407) ←
Hammonia hammondi (37729)
Bartonella henselae (1471)
Feline immunodeficiency virus (1391) ←
Yersinia pestis PY-47 (1070)
All other taxa (6021)
More...

Find related data

Database: Select

Find items

Search details

"Felis catus"[Organism] OR Felis catus[All Fields]

Search See more...

A better way to get the cat genome would be via the Genome database

← → C Secure <https://www.ncbi.nlm.nih.gov/genome/?term=felis+catus>

NCBI Resources How To

Genome

Genome felis catus[orgn]

Create alert Limits Advanced

Felis catus (domestic cat)
Reference genome: [Felis catus \(assembly Felis_catus_8.0\)](#)
Download sequences in FASTA format for [genome](#), [transcript](#), [protein](#)
Download genome annotation in [GFF](#), [GenBank](#) or [tabular](#) format
[BLAST](#) against Felis catus [genome](#)
All 2 genomes for species:
[Browse the list](#)
[Download sequence and annotation from RefSeq or GenBank](#)

Display Settings: Overview Send to: ID: 78

[Organism Overview](#) ; [Genome Assembly and Annotation report \[2\]](#) ; [Organelle Annotation Report \[1\]](#)

 **Felis catus (domestic cat)**
domestic cat

Lineage: [Eukaryota\[2334\]](#); [Metazoa\[779\]](#); [Chordata\[332\]](#); [Craniata\[324\]](#); [Vertebrata\[324\]](#); [Euteleostomi\[319\]](#); [Mammalia\[136\]](#); [Eutheria\[131\]](#); [Laurasiatheria\[61\]](#); [Carnivora\[13\]](#); [Feliformia\[4\]](#); [Felidae\[4\]](#); [Felineae\[1\]](#); [Felis\[1\]](#); [Felis catus\[1\]](#)

Felis catus, the domestic cat, provides several valuable models for infectious disease, including a model for human AIDS. With a large number of recognized breeds, the cat is also a valuable resource for studying phenotypic diversity and evolution. The cat genome will further facilitate research in human medicine as some rare diseases that occur [More...](#)

There are actually 2 cat genome assemblies in NCBI

← → C Secure | <https://www.ncbi.nlm.nih.gov/genome/?term=felis+catus>

NCBI Resources How To

Genome Genome felis catus[orgn]
Create alert Limits Advanced

Felis catus (domestic cat)
Reference genome: [Felis catus \(assembly Felis_catus_8.0\)](#)
Download sequences in FASTA format for [genome](#), [transcript](#), [protein](#)
Download genome annotation in [GFF](#), [GenBank](#) or [tabular](#) format
BLAST against Felis catus [genome](#)
All 2 genomes for species: 
[Browse the list](#)
[Download sequence and annotation from RefSeq or GenBank](#)

Display Settings: Overview Send to: ID: 78

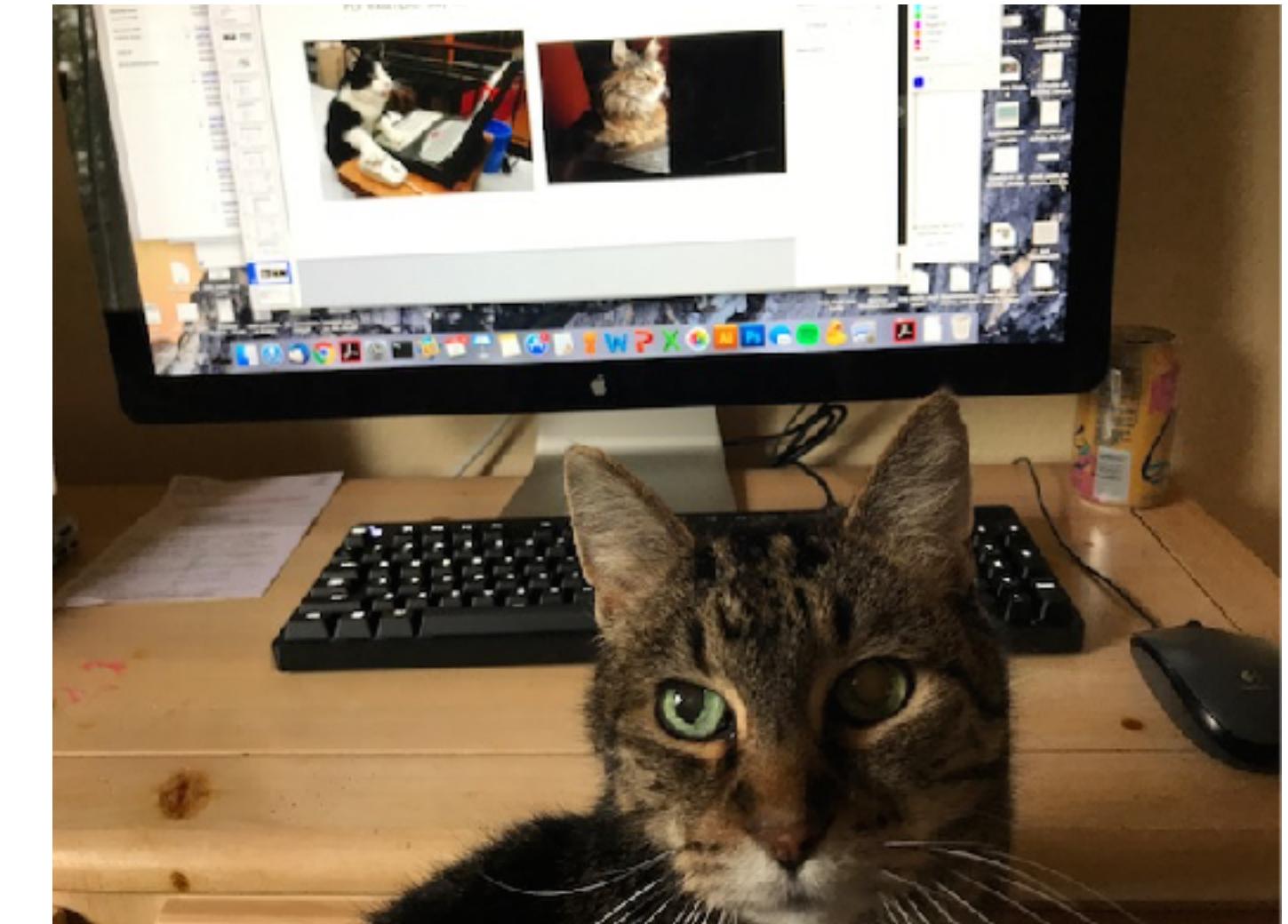
[Organism Overview](#) ; [Genome Assembly and Annotation report \[2\]](#) ; [Organelle Annotation Report \[1\]](#)

 **Felis catus (domestic cat)**
domestic cat

Lineage: [Eukaryota](#)[2334]; [Metazoa](#)[779]; [Chordata](#)[332]; [Craniata](#)[324]; [Vertebrata](#)[324]; [Euteleostomi](#)[319]; [Mammalia](#)[136]; [Eutheria](#)[131]; [Laurasiatheria](#)[61]; [Carnivora](#)[13]; [Feliformia](#)[4]; [Felidae](#)[4]; [Felineae](#)[1]; [Felis](#)[1]; [Felis catus](#)[1]

Felis catus, the domestic cat, provides several valuable models for infectious disease, including a model for human AIDS. With a large number of recognized breeds, the cat is also a valuable resource for studying phenotypic diversity and evolution. The cat genome will further facilitate research in human medicine as some rare diseases that occur [More...](#)

In reality, there are at least as many cat genomes as there are cats



Kirby

You could also get at the cat genome via the Taxonomy database

Secure | https://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi

NCBI Taxonomy Browser

Search for: Felis catus as complete name lock Go

Display 3 levels using filter: none

Felis catus

Taxonomy ID: 9685
Genbank common name: domestic cat
Inherited blast name: carnivores
Rank: species
Genetic code: Translation table 1 (Standard)
Mitochondrial genetic code: Translation table 2 (Vertebrate Mitochondrial)
Other names:
synonym: **Felis silvestris catus**
synonym: **Felis domesticus**
common name: **cats**
common name: **cat**
includes: **Korat cats**
authority: **Felis catus Linnaeus, 1758**

[Lineage \(full\)](#)
[cellular organisms](#); [Eukaryota](#); [Opisthokonta](#); [Metazoa](#); [Eumetazoa](#); [Bilateria](#); [Deuterostomia](#); [Chordata](#); [Craniata](#); [Vertebrata](#); [Gnathostomata](#); [Teleostomi](#); [Euteleostomi](#); [Sarcopterygii](#); [Dipnotetrapodomorpha](#); [Tetrapoda](#); [Amniota](#); [Mammalia](#); [Theria](#); [Eutheria](#); [Boreoeutheria](#); [Laurasiatheria](#); [Carnivora](#); [Feliformia](#); [Felidae](#); [Felinae](#); [Felis](#)

Entrez records	
Database name	Direct links
Nucleotide	355,415
Nucleotide EST	919
Nucleotide GSS	3,107
Protein	55,602
Structure	10
Genome	1
Popset	172
SNP	3,611,275
GEO Datasets	103
PubMed Central	1,706
Gene	39,815
SRA Experiments	406
Probe	2,877
Assembly	4
Bio Project	39
Bio Sample	396
Bio Systems	495
Clone DB	239,767
PubChem BioAssay	1,097
Protein Clusters	12
Taxonomy	1

You can go up the taxonomic tree in the Taxonomy db

Secure | https://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi

NCBI Taxonomy Browser

Search for: Felis catus as complete name lock Go

Display 3 levels using filter: none

Felis catus

Taxonomy ID: 9685
Genbank common name: domestic cat
Inherited blast name: carnivores
Rank: species
Genetic code: Translation table 1 (Standard)
Mitochondrial genetic code: Translation table 2 (Vertebrate Mitochondrial)
Other names:
synonym: **Felis silvestris catus**
synonym: **Felis domesticus**
common name: cats
common name: cat
includes: Korat cats
authority: **Felis catus Linnaeus, 1758**

[Lineage \(full\)](#)
[cellular organisms](#); [Eukaryota](#); [Opisthokonta](#); [Metazoa](#); [Eumetazoa](#); [Bilateria](#); [Deuterostomia](#); [Chordata](#); [Craniata](#); [Vertebrata](#); [Gnathostomata](#); [Teleostomi](#); [Euteleostomi](#); [Sarcopterygii](#); [Dipnotetrapodomorpha](#); [Tetrapoda](#); [Amniota](#); [Mammalia](#); [Theria](#); [Eutheria](#); [Boreoeutheria](#); [Laurasiatheria](#); [Carnivora](#); [Feliformia](#); [Felidae](#); [Felinae](#); [Felis](#)

Entrez records

Database name	Direct links
Nucleotide	355,415
Nucleotide EST	919
Nucleotide GSS	3,107
Protein	55,602
Structure	10
Genome	1
Popset	172
SNP	3,611,275
GEO Datasets	103
PubMed Central	1,706
Gene	39,815
SRA Experiments	406
Probe	2,877
Assembly	4
Bio Project	39
Bio Sample	396
Bio Systems	495
Clone DB	239,767
PubChem BioAssay	1,097
Protein Clusters	12
Taxonomy	1

You can go up the taxonomic tree in the Taxonomy db

NCBI Taxonomy Browser

Search for: as complete name lock Go

Display 3 levels using filter: none

Nucleotide Nucleotide EST Nucleotide GSS Protein Structure Genome Popset SNP
 Domains GEO Datasets UniGene PubMed Central Gene HomoloGene SRA Experiments Probe
 Assembly LinkOut BLAST TRACE Host Viral Host Bio Project Bio Sample
 Bio Systems Clone DB dbVar GEO Profiles PubChem BioAssay Protein Clusters

Lineage (full): root; cellular organisms; Eukaryota; Opisthokonta; Metazoa; Eumetazoa; Bilateria; Deuterostomia; Chordata; Craniata; Vertebrata; Amniota; Mammalia; Theria; Eutheria; Boreoeutheria; Laurasiatheria; Carnivora; Feliformia

- **Felidae** (cat family) [36](#) Click on organism name to get more information.
 - [Acinonychiae](#) [1](#)
 - [Acinonyx](#) [1](#)
 - [Acinonyx jubatus](#) (cheetah) [1](#) ←
 - [Felinae](#) [28](#)
 - [Caracal](#) [1](#)
 - [Caracal caracal](#) [1](#)
 - [Catopuma](#) [2](#)
 - [Catopuma badia](#) (bay cat) [1](#)
 - [Catopuma temminckii](#) (Asiatic golden cat) [1](#)
 - [Felis](#) [5](#)
 - [Felis catus](#) (domestic cat) [1](#)
 - [Felis chaus](#) (jungle cat) [1](#)
 - [Felis margarita](#) (sand cat) [1](#)
 - [Felis nigripes](#) (black-footed cat) [1](#)
 - [Felis silvestris](#) (wild cat) [1](#)
 - [Felis sp.](#)
 - [Felis sp. NG192](#)
 - [Leopardus](#) [7](#)
 - [Leopardus colocolo](#) (Colocolo) [1](#)
 - [Leopardus geoffroyi](#) (Geoffroy's cat) [1](#)
 - [Leopardus guigna](#) (Kodkod) [1](#)
 - [Leopardus guttulus](#) (Southern oncilla)
 - [Leopardus jacobita](#) (Andean mountain cat) [1](#)
 - [Leopardus pardalis](#) (ocelot) [1](#)
 - [Leopardus tigrinus](#) (little spotted cat) [1](#)
 - [Leopardus wiedii](#) (margay) [1](#)
 - [Leptailurus](#) [1](#)
 - [Leptailurus serval](#) (serval) [1](#)
 - [Leptailurus serval x Caracal caracal](#)
 - [Lynx](#) [4](#)
 - [Lynx canadensis](#) (Canada lynx) [1](#)
 - [Lynx lynx](#) (Eurasian lynx) [1](#)
 - [Lynx pardinus](#) (Spanish lynx) [1](#)
 - [Lynx rufus](#) (bobcat) [1](#)
 - [environmental samples](#)
- [Otocolobus](#) [1](#)

You can go up the taxonomic tree in the Taxonomy db

NCBI Resources How To

Genome Genome txid32536[Organism:exp] | Create alert Limits Advanced

Acinonyx jubatus (cheetah)
Representative genome: [Acinonyx jubatus \(assembly aciJub1\)](#)
Download sequences in FASTA format for [genome](#), [transcript](#), [protein](#)
Download genome annotation in [GFF](#), [GenBank](#) or [tabular](#) format
[BLAST against Acinonyx jubatus genome](#)

Display Settings: Overview Send to: ID: 14623

[Organism Overview](#) ; [Organelle Annotation Report \[1\]](#)

Acinonyx jubatus (cheetah)

Acinonyx jubatus overview

Lineage: [Eukaryota](#)[2334]; [Metazoa](#)[779]; [Chordata](#)[332]; [Craniata](#)[324]; [Vertebrata](#)[324]; [Euteleostomi](#)[319]; [Mammalia](#)[136]; [Eutheria](#)[131]; [Laurasiatheria](#)[61]; [Carnivora](#)[13]; [Feliformia](#)[4]; [Felidae](#)[4]; [Acinonychinae](#)[1]; [Acinonyx](#)[1]; [Acinonyx jubatus](#)[1]

Summary

Submitter: Saint Petersburg State University
Assembly level: Scaffold
Environment: OptimumTemperature: C
Assembly: GCA_001443585.1 aciJub1 scaffolds: 14,383 contigs: 170,171 N50: 35,120 L50: 19,059
BioProjects: PRJNA305718, PRJNA297632
Whole Genome Shotgun (WGS): INSDC: LLWD00000000.1
Statistics: total length (Mb): 2372.55
protein count: 27284
GC%: 41.4
NCBI Annotation Release: 100

Publications

You need not rely on your browser to download data

← → ⌂ **Secure** | <https://www.ncbi.nlm.nih.gov/genome/?term=felis+catus>

NCBI Resources How To

Genome [Genome](#) [felis catus\[orgn\]](#) Create alert Limits Advanced

Felis catus (domestic cat)
Reference genome: [Felis catus \(assembly Felis_catus_8.0\)](#)
Download sequences in FASTA format for genome, transcript, protein
Download genome annotation in [GFF](#), [GenBank](#) or tabular format
BLAST against [Felis catus genome](#)

All 2 genomes for species:
[Browse the list](#)
Download sequence and annotation from [RefSeq](#) or [GenBank](#)

FTP links ←

Display Settings: Overview Send to: ID: 78

Organism Overview ; Genome Assembly and Annotation report [2] ; Organelle Annotation Report [1]

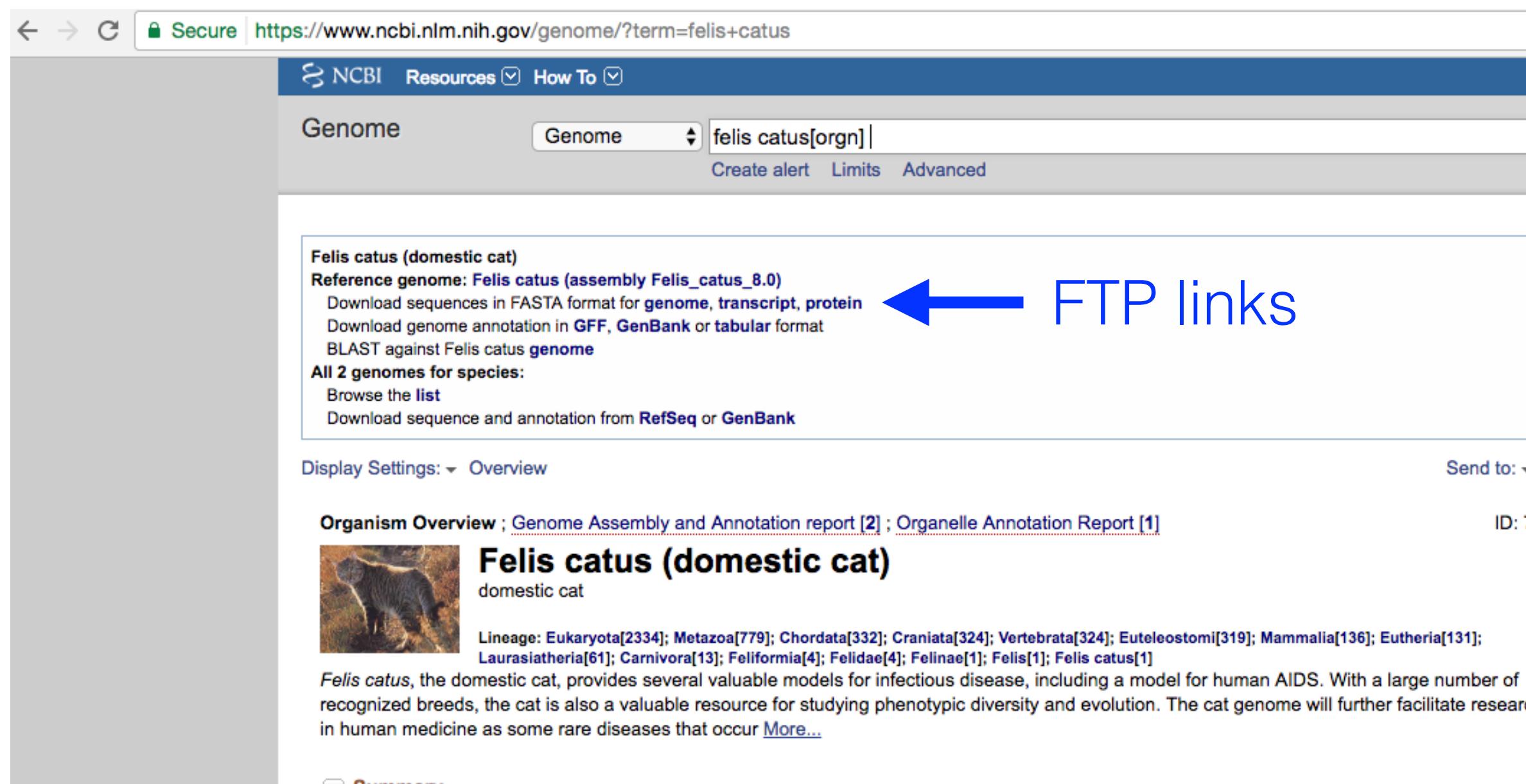
Felis catus (domestic cat)
domestic cat

 Lineage: Eukaryota[2334]; Metazoa[779]; Chordata[332]; Craniata[324]; Vertebrata[324]; Euteleostomi[319]; Mammalia[136]; Eutheria[131]; Laurasiatheria[61]; Carnivora[13]; Feliformia[4]; Felidae[4]; Felinae[1]; Felis[1]; Felis catus[1]

Felis catus, the domestic cat, provides several valuable models for infectious disease, including a model for human AIDS. With a large number of recognized breeds, the cat is also a valuable resource for studying phenotypic diversity and evolution. The cat genome will further facilitate research in human medicine as some rare diseases that occur [More...](#)

You can download data from the command line

This is often useful when you're working on a server.



Secure https://www.ncbi.nlm.nih.gov/genome/?term=felis+catus

NCBI Resources How To

Genome

Genome felis catus[orgn]

Create alert Limits Advanced

Felis catus (domestic cat)

Reference genome: **Felis catus (assembly Felis_catus_8.0)**

Download sequences in FASTA format for **genome, transcript, protein**
Download genome annotation in **GFF, GenBank or tabular format**
BLAST against Felis catus genome

All 2 genomes for species:

Browse the list
Download sequence and annotation from **RefSeq or GenBank**

Display Settings: Overview

Send to:

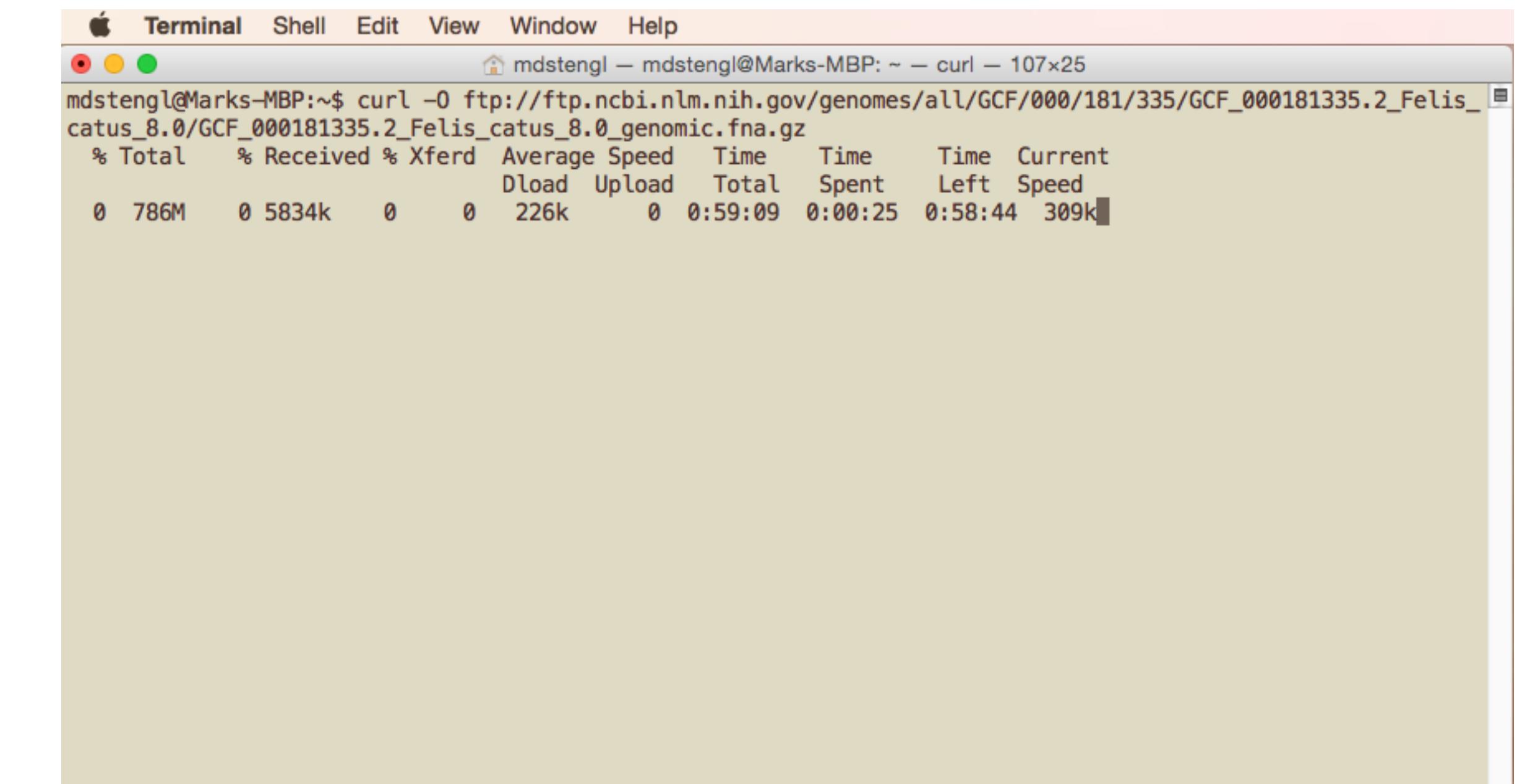
Organism Overview ; Genome Assembly and Annotation report [2] ; Organelle Annotation Report [1]

ID: 78

Felis catus (domestic cat)
domestic cat

Lineage: Eukaryota[2334]; Metazoa[779]; Chordata[332]; Craniata[324]; Vertebrata[324]; Euteleostomi[319]; Mammalia[136]; Eutheria[131]; Laurasiatheria[61]; Carnivora[13]; Feliformia[4]; Felidae[4]; Felinae[1]; Felis[1]; Felis catus[1]

Felis catus, the domestic cat, provides several valuable models for infectious disease, including a model for human AIDS. With a large number of recognized breeds, the cat is also a valuable resource for studying phenotypic diversity and evolution. The cat genome will further facilitate research in human medicine as some rare diseases that occur [More...](#)



```
Terminal Shell Edit View Window Help
mdstengl@Marks-MBP:~$ curl -O ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCF/000/181/335/GCF_000181335.2_Felis_catus_8.0/GCF_000181335.2_Felis_catus_8.0_genomic.fna.gz
% Total    % Received % Xferd  Average Speed   Time     Time     Time  Current
          Dload  Upload Total   Spent    Left Speed
0  786M     0  5834k    0      0  226k      0  0:59:09  0:00:25  0:58:44  309k
```

curl is a file transfer utility built into Linux, MacOS

similar utilities exist for Windows

GUI-based software for file transfer

Cyberduck



Screenshot of the Cyberduck application interface. The title bar shows "ftp.ncbi.nlm.nih.gov - FTP" and "Unregistered". The main window displays a file list with columns for "Filename", "Size", and "Modified". The list includes various directories and files such as "genomes", "bioproject", "biosample", "snp", "repository", "pubchem", "pathogen", "hmm", "geo", "pubmed", "refseq", "sra", "ncbi-asn1", "genbank", "eqtl", "dbgap", "pub", "sequin", "variation", and "nist-immsa". The total count at the bottom is "46 Files".

Filename	Size	Modified
genomes	32.8 KB	Today, 9:23 AM
bioproject	4.1 KB	Today, 9:10 AM
biosample	4.1 KB	Today, 2:39 AM
snp	4.1 KB	Yesterday, 8:48 PM
repository	4.1 KB	Yesterday, 8:48 PM
pubchem	4.1 KB	Yesterday, 8:48 PM
pathogen	4.1 KB	Yesterday, 8:48 PM
hmm	4.1 KB	Yesterday, 1:20 PM
geo	0 B	Yesterday, 6:22 AM
pubmed	4.1 KB	5/24/17, 2:05 AM
refseq	4.1 KB	5/15/17, 11:38 AM
sra	4.1 KB	5/4/17, 11:38 AM
ncbi-asn1	73.7 KB	4/19/17, 9:20 PM
genbank	94.2 KB	4/19/17, 9:07 PM
eqtl	4.1 KB	4/15/17, 4:14 PM
dbgap	4.1 KB	4/3/17, 1:34 PM
pub	8.2 KB	3/20/17, 10:58 AM
sequin	4.1 KB	2/22/17, 6:50 AM
variation	4.1 KB	8/9/16, 8:36 AM
nist-immsa	4.1 KB	6/30/16, 12:19 PM

<ftp://ftp.ncbi.nlm.nih.gov/>



You can search and download from NCBI databases directly in Geneious

Geneious 10.0.7

Back Forward BLAST Workflows Align/Assemble Tree Primers Cloning Back Up Support Help

Sources

Local (84317, 18982 unread) Shared Databases Operations NCBI Gene Genome (20) Nucleotide NCBI Genome database (32,686 records) Protein PubMed SNP Structure Taxonomy UniProt

Download completed

Felis catus

Sequence Length	Name	Description
94,101,111	NC_018734	Felis catus isolate Cinnamon breed Abyssinian chromosome D3, Felis_catus_8.0, whole genome shotgun sequence
88,096,124	NC_018733	Felis catus isolate Cinnamon breed Abyssinian chromosome D2, Felis_catus_8.0, whole genome shotgun sequence
83,953,389	NC_018740	Felis catus isolate Cinnamon breed Abyssinian chromosome F2, Felis_catus_8.0, whole genome shotgun sequence
70,119,229	NC_018739	Felis catus isolate Cinnamon breed Abyssinian chromosome F1, Felis_catus_8.0, whole genome shotgun sequence
61,960,243	NC_018737	Felis catus isolate Cinnamon breed Abyssinian chromosome E2, Felis_catus_8.0, whole genome shotgun sequence
61,081,816	NC_018736	Felis catus isolate Cinnamon breed Abyssinian chromosome E1, Felis_catus_8.0, whole genome shotgun sequence
41,224,383	NC_018738	Felis catus isolate Cinnamon breed Abyssinian chromosome E3, Felis_catus_8.0, whole genome shotgun sequence
17,009	NC_001700	Felis catus mitochondrial, complete genome

Sequence View Annotations Dotplot (Self) Text View Fasta Nucleotide View Summary XML Linear

Extract R.C. Translate Add/Edit Annotation Allow Editing Annotate & Predict Save

NC_001700
17,009 bp

contro region; CR

CYTB CDS

ND6 CDS

ND5 CDS

ND4 CDS

ND4L CDS

ND3 CDS

ND2 CDS

ND1 CDS

COX1 gene

COX2 gene

ATP6 gene

ATP8 gene

ND4L gene

ND3 gene

COX3 gene

AIP6 CDS

ATP8 CDS

control region; CR

cyt b gene

ND6 gene

ND5 gene

ND4 gene

ND2 gene

ND1 gene

I-
rRNA

5'-
rRNA

RH136367

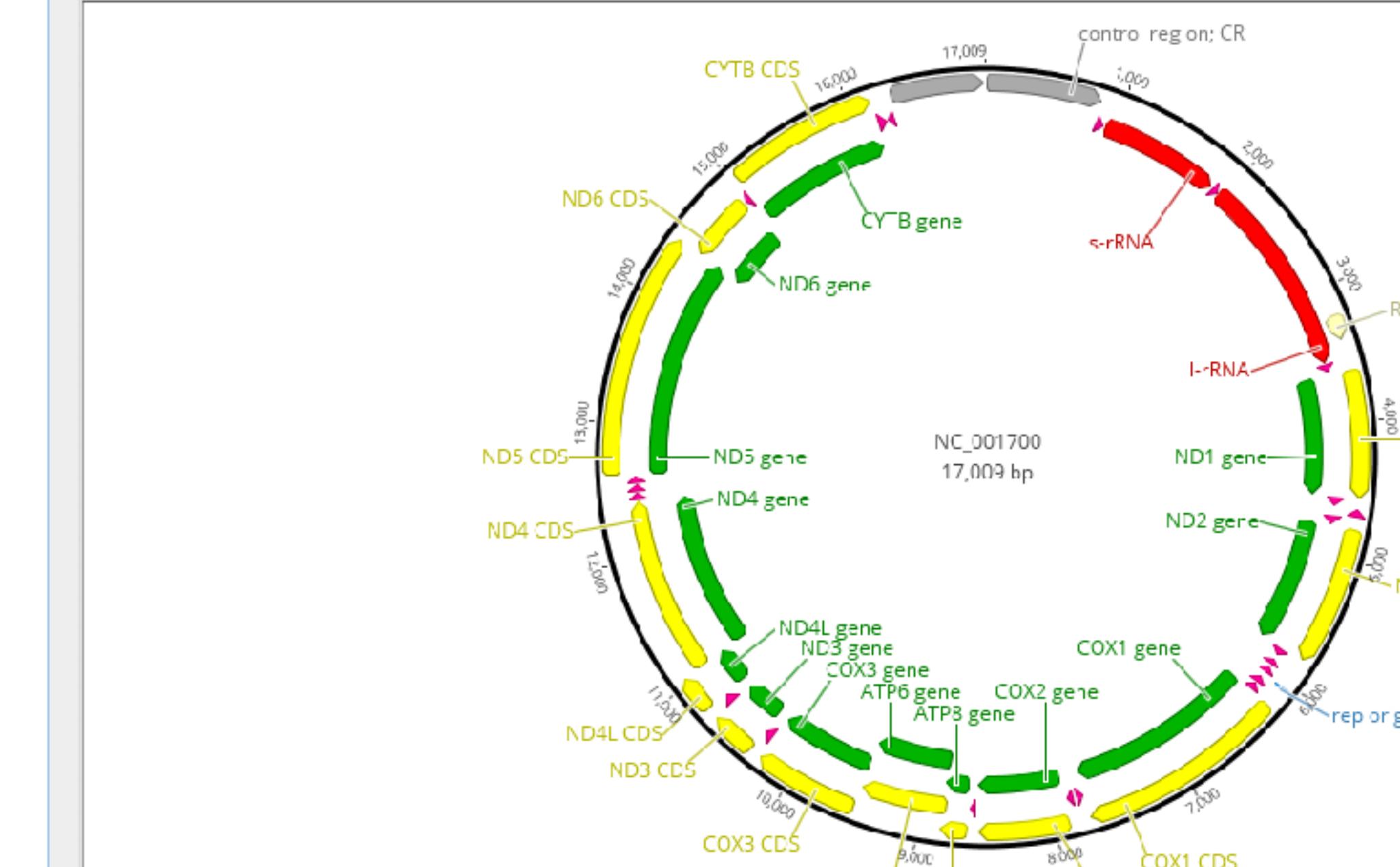
rep origin

Alt click on a sequence position or annotation, or select a region to zoom in. Alt-shift click to zoom out.

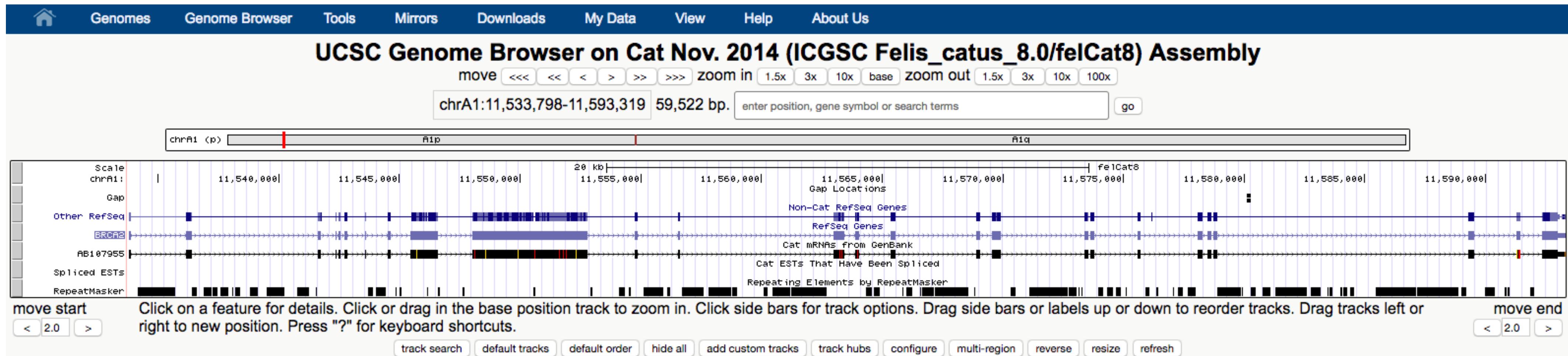
Updating search index: 2 items

+ -

Using 245 / 7996 MB memory



Genome browsers, like Ensembl and UCSC, offer additional functionality



Finally, there's absolutely nothing wrong with using Google

cat genome

All Shopping Images News Videos More Settings Tools

About 11,500,000 results (0.60 seconds)

Specifically, the signatures of selection in the domestic cat genome are linked to genes associated with gene knockout models affecting memory, fear-conditioning behavior, and stimulus-reward learning, and potentially point to the processes by which cats became domesticated. Dec 2, 2014

Comparative analysis of the domestic cat genome reveals genetic ...
www.pnas.org/content/111/48/17230.abstract

The cat genome - NCBI
https://www.ncbi.nlm.nih.gov/genome?term=felis%20catus
1. Felis catus domestic cat Kingdom: Eukaryota Chromosomes: 19 Organelles: 1 Genome ID: 78.

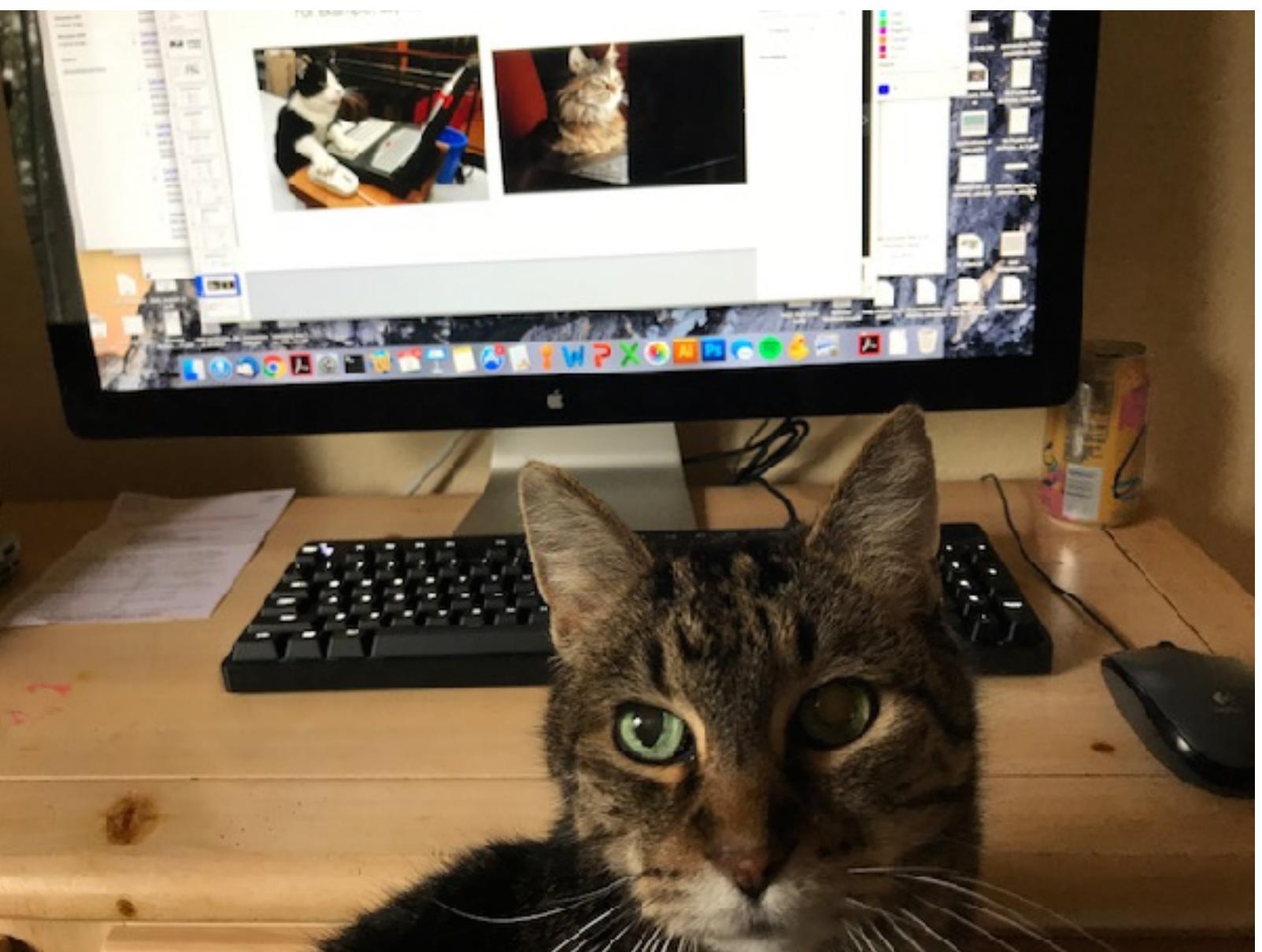
Comparative analysis of the domestic cat genome reveals genetic ...
www.pnas.org/content/111/48/17230.abstract ▾
by MJ Montague - 2014 - Cited by 67 - Related articles
Dec 2, 2014 - Specifically, the signatures of selection in the domestic cat genome are linked to genes associated with gene knockout models affecting memory, fear-conditioning behavior, and stimulus-reward learning, and potentially point to the processes by which cats became domesticated.

Felis catus - Ensembl genome browser 89
www.ensembl.org/Felis_catus/Info/Index ▾
What's New in Cat release 89. Microarray Probe Mapping ... Genome assembly: Felis_catus_6.2 (GCA_000181335.2) ... assemblies. CAT (Ensembl release 67).

'I can haz genomes': cats claw their way into genetics : Nature News ...
www.nature.com/news/i-can-haz-genomes-cats-claw-their-way-into-genetics-1.16708 ▾
Jan 14, 2015 - The cat genome is out of the bag, and has already helped to pinpoint a gene involved in kidney disease. ... The work will benefit both humans and felines, the researchers say, by mapping the mutations underlying conditions that afflict the two species, such as kidney disease.

The Cat Genome Sequence - Lyons Feline & Comparative Genetics
felinegenetics.missouri.edu/feline-genome-project-2/the-cat-genome-sequence ▾
The Cat Genome Sequence. NCBI – Cat Genome Resources http://www.ncbi.nlm.nih.gov/projects/genome/guide/cat/ · Science Daily – Domestic Cat Genome ...

Questions?



Kirby in 2000, wondering where his GenBank CDROMs are